electronic reprint

Journal of
Applied Crystallography
ISSN 0021-8898

Crystallization of protein–protein complexes

Sergei Radaev and Peter D. Sun

Journal of
Applied Crystallography
ISSN 0021-8898

Received 1 August 2002
Accepted 5 August 2002

Crystallization of protein–protein complexes

Sergei Radaev and Peter D. Sun*

Structural Immunology Section, Laboratory of Immunogenetics, National Institute of Allergy and Infectious Diseases, National Institutes of Health, 12441 Parklawn Drive, Rockville, Maryland 20852, USA. Correspondence e-mail: psun@nih.gov

Crystallizing protein–protein complexes remains a rate-limiting step in their structure characterization. Crystallization conditions for the known protein–protein complexes have been surveyed in both the Protein Data Bank and the BMCD database. Compared with non-complexed proteins, crystallization conditions for protein–protein complexes are less diverse and heavily favor (71% *versus* 27%) polyethylene glycols (PEG) rather than ammonium sulfate or other high-salt crystallization conditions. The results suggest that the stability of protein complexes limits their available crystallization configuration space. Based on the survey, a set of sparse-matrix screen conditions was designed.

1. Introduction

In the age of structural genomics, crystallization remains a major rate-limiting step in the process of macromolecular structure determination (McPherson, 1999; Bergfors, 1999). The problem arises from the need to sample a seemingly unlimited crystallization configuration space with often a very limited amount of protein available. There have been several attempts to develop a rational approach for macromolecular crystallization. Some of them were based solely on a mathematical approach, such as an incomplete factorial procedure (Carter & Carter, 1979; Carter, 1990) in which all conditions in a multi-dimensional crystallization phase diagram are considered as equally probable and a small number of experimental conditions are chosen uniformly throughout the diagram to provide an even sampling of its entire volume. An incomplete factorial approach was applied to develop a sparse-matrix kit available from Hampton Research (Jancarik & Kim, 1991). This approach is based on a random search for crystallization conditions and does not take into account any differences between various classes of macromolecules, which may significantly narrow the search area. A recent analysis of the crystallization conditions deposited in the BMCD (Gilliland & Bickham, 1990) led to the conclusion that different classes of macromolecules show systematic biases in their pattern of successful crystallization (Hennessy *et al.*, 2000). Increasingly, studies of biological functions require crystallization of protein–protein complexes. Our experience in crystallizing receptor–ligand complexes (Boyington *et al.*, 2000; Radaev, Motyka *et al.*, 2001; Radaev, Rostro *et al.*, 2001) indicates the existence of preferable conditions for crystallization of protein–protein complexes. To generalize the approach further, we surveyed the known crystallization conditions for all protein–protein complexes available to date. Based on this survey, a 'sparse-matrix' screen was designed to reflect the most probable conditions for crystallizing protein–protein complexes.

2. Material and methods

To prepare the crystallization database for protein–protein complexes, a search of the Protein Data Bank (PDB) was performed using the keyword 'complex' with DNA–protein and RNA–protein complexes excluded. A manual examination of 900 hits was then carried out to exclude protein–peptide complexes and to reduce redundant entries for the same protein–protein complex to one entry for each space group and crystallization condition. A size of 40 amino acids was used as a selection criterion to distinguish short peptides from proteins in protein–protein complexes; antibody–antigen complexes with small haptens were also excluded. In addition, a systematic examination of the BMCD crystallization database was carried out to identify valid candidates using the same criteria. This search resulted in about 200 unique entries in the protein–protein complex crystallization database. Among them, there are 38 entries of antibody (Fab)–antigen complexes, 35 cytokine receptor–ligand complexes, 25 enzyme–inhibitor complexes, 27 immuno-receptor and ligand complexes, 34 complexes of signal transduction proteins, 19 oligomeric protein complexes, such as proteasome, cytochrome C oxidase and photoreaction centers; and 22 miscellaneous protein–protein complexes. A random sample of 100 proteins, excluding protein–protein complexes, was selected from the BMCD database to form a control set of non-complexed proteins.

3. Results and discussion

3.1. Precipitation reagents

To analyze statistical preference in precipitants used for crystallizing protein–protein complexes, the crystallization conditions were grouped into four categories: (i) PEG (all molecular weight combined); (ii) $(NH_4)_2SO_4$; (iii) salts other

Table 1
Protein complex crystallization screening kit.

	Precipitant	Salt	Buffer†
1	20% PEG 400	0.1 *M* $CaAc_2$	HEPES/7.0
2	20% MPEG 550		Tris/8.0
3	20% PEG 1000		MES/6.5
4	20% PEG 1500		MES/6.5
5	15% PEG 2000	0.15 *M* NaCl	Tris/8.5
6	20% MPEG 2000		Citrate/6.5
7	20% MPEG 2000	0.2 *M* NaCl	HEPES/7.0
8	20% MPEG 2000	0.1 *M* Li_2SO_4	MES/6.0
9	15% PEG 3350	0.1 *M* $(NH_4)_2SO_4$	HEPES/7.5
10	20% PEG 3350	0.1 *M* NaCl	Tris/8.0
11	20% PEG 3350	0.1 *M* $MgCl_2$	MOPS/7.0
12	20% PEG 3350		Cacodylate/6.5
13	8% PEG 4000		MES/6.0
14	12% PEG 4000		Citrate/5.5
15	12% PEG 4000	0.1 *M* $(NH_4)_2SO_4$	Citrate/5.5
16	15% PEG 4000	0.1 *M* $MgCl_2$	Cacodylate/6.0
17	15% PEG 4000	0.1 *M* $CaCl_2$	MES/6.5
18	15% PEG 4000	10% 2-propanol	HEPES/7.0
19	15% PEG 4000	0.1 *M* NaCl	Tris/8.0
20	20% PEG 4000		HEPES/7.0
21	10% MPEG 5000	10% 2-propanol	Cacodylate/6.0
22	20% MPEG 5000		HEPES/7.5
23	8% PEG 6000		MES/6.0
24	15% PEG 6000		NaAc/5.0
25	10% PEG 6000	0.1 *M* NaCl	HEPES/7.0
26	10% PEG 6000	0.1 *M* $CaAc_2$	HEPES/7.5
27	15% PEG 6000		Tris/8.0
28	8% PEG 8000		MES/6.0
29	12% PEG 8000		Citrate/5.5
30	12% PEG 8000	0.1 *M* NaCl	Cacodylate/6.0
31	12% PEG 8000	0.1 *M* $MgAc_2$	MES/6.5
32	12% PEG 8000	0.1 *M* KCl	HEPES/7.0
33	12% PEG 8000	0.1 *M* $(NH_4)_2SO_4$	HEPES/7.5
34	12% PEG 8000		Tris/8.0
35	12% PEG 8000	0.1 *M* $MgAc_2$	Tris/8.5
36	8% PEG 20000		MES/6.0
37	12% PEG 20000		Tris/8.0
38	12% PEG 20000		HEPES/7.0
39	12% PEG 20000	0.1 *M* $CaCl_2$	Cacodylate/6.5
40	1 *M* $(NH_4)_2SO_4$		MES/6.0
41	1.5 *M* $(NH_4)_2SO_4$		Citrate/5.5
42	1.6 *M* $(NH_4)_2SO_4$		Cacodylate/6.5
43	1.6 *M* $(NH_4)_2SO_4$		HEPES/7.5
44	1.8 *M* $(NH_4)_2SO_4$		Tris/8.5
45	1 *M* NaCl		Tris/8.0
46	1 *M* K-,Na-tartrate		HEPES/7.0
47	1.3 *M* Li_2SO_4		Tris/8.0
48	1.5 *M* Na_2KPO_4		Tris/8.0

† All buffers are at 0.1 *M* concentration.

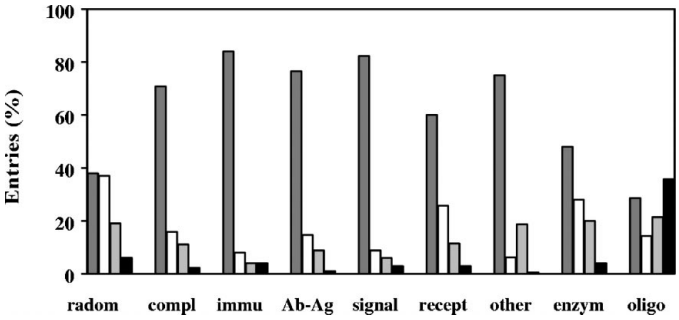

Figure 1
Percentage of different types of precipitants used in the crystallization of protein–protein complexes: PEG, $(NH_4)_2SO_4$, other salts and organic solvents (including 2-propanol, MPD and ethanol) are shown as medium-gray, white, light-gray and black bars, respectively. The categories 'random', 'compl', 'immu', 'Ab-Ag', 'signal', 'recept', 'enzyme', 'oligo', and 'other' are for random samples, all protein–protein complexes included in this survey, immune complexes, antibody–antigen complexes, signal transduction complexes, receptor and ligand complexes, enzyme-related complexes, oligomeric protein complexes and other miscellaneous protein–protein complexes, respectively. The frequencies are normalized to 100% within each category. Multiple entries of the same protein–protein complex are reduced to unique entries based on space group and crystallization conditions.

than $(NH_4)_2SO_4$; and (iv) organic solvents (2-propanol, MPD, ethanol) (Fig. 1).

Overall, excluding the oligomeric protein complexes, the survey shows the highest success rate for PEG in protein–protein complex crystallization, with 71% of cases crystallized from PEG, 16% from ammonium sulfate, 11% from other salts, and 2% from organic solvents. In contrast, among randomly chosen non-complexed proteins, the success rate for PEG is almost equal to that for $(NH_4)_2SO_4$; the distribution is 38, 37, 19 and 6% for PEG, $(NH_4)_2SO_4$, other salts and organic solvents, respectively. The preference for using PEG as precipitant is also obvious in each sub-family of proteins, such as antibody–antigen complexes, immune complexes, *etc*. (Fig. 1). The trend in favor of PEG in crystallization conditions disappears, however, for the family of oligomeric protein complexes, such as photoreaction centers, ribosomal subunits, and nucleosomes. When the individual crystallization conditions are examined, there appears to be a correlation between the choice of precipitant and the stability of the complexes. Lower-affinity protein–protein complexes tend to favor PEG conditions, whereas for higher-affinity complexes, $(NH_4)_2SO_4$ also becomes a very successful precipitant. For example, immune complexes KIR2DL2/HLA-Cw3 and IgG1-Fc/FcγRIII with $K_D = 10^{-5}$–10^{-6} *M* could be crystallized from PEG only (Boyington *et al.*, 2000; Radaev, Motyka *et al.*, 2001), whereas crystals of the high-affinity ($K_D = 10^{-10}$ *M*) immune complex IgE-Fc/FcεRI were grown from 1.4–1.6 *M* $(NH_4)_2SO_4$ (Garman *et al.*, 2000).

3.2. Polyethylene glycols

Among all the crystallizations of protein–protein complexes in which polyethylene glycols are used as precipitants, the most popular ones employ PEG with molecular weights between 3000 and 8000. Together, they count for 80% of all the PEG conditions in this survey. A small number of cases (about 10%) are crystallized from PEG MME (polyethylene glycol monomethyl ether) and MPEG (methoxy polyethylene glycol).

The PEG concentration in protein complex crystallizations ranges from 5 to 30%, similar to that of normal soluble proteins. The majority, however, are between 10 and 20% of PEG. In general, the optimum concentration of PEG required to crystallize individual proteins depends mostly on the solubility and concentration of a given protein and can vary drastically from protein to protein. In many cases, the solubility of a protein complex is less than that of its components, thus requiring a lower concentration of PEG for crystallization.

3.3. Buffer, pH and salts

Extreme pH conditions are generally not favorable for protein crystallization because of their destabilizing effect on protein stability. However, many soluble proteins are stable and can be crystallized, for example, from pH 2 to 5. As a result, low pH conditions are usually not excluded from initial crystal screen experiments. For example, one out of seven conditions in the Hampton Research crystal screen kit has a pH of 4.6. In the case of a protein complex crystallization, the available pH range for crystallization is expected to depend on the stability of a protein–protein complex. Our current survey shows that the majority of the protein–protein complexes are crystallized between pH 6.0 and 8.5. There are no reported complexes that crystallized below pH 4.5 and six complexes (3% of the total survey) crystallized between pH 4.5 and 4.9. The percentage of protein complexes that crystallized at low pH values is much less than that represented in the Hampton Research crystal screen kit. Within the mild pH range, there is no clear preference for buffer type in crystallizing protein–protein complexes. In fact, the use of various buffers is quite evenly distributed, suggesting that a screening of pH and buffer type is necessary.

Most protein–protein complexes crystallized using PEG as a precipitant contain salts at moderate concentrations, of generally not more than 300 m*M*. The most commonly used salts are NaCl, $MgCl_2$, $Mg(CH_3COO)_2$, $CaCl_2$, $(NH_4)_2SO_4$, Li_2SO_4, $NH_4(CH_3COO)$, KCl, $Na(CH_3COO)$, $Ca(CH_3COO)_2$, and Na-/K-tartrate. When ammonium sulfate is used as a major precipitant, its optimal concentration ranges from 1.0 to 2.0 *M*.

3.4. Protein–protein complex crystallization screening kit

As described here, several protein–protein complex crystallization parameters, such as precipitant, pH and salt have distinct distributions compared with general soluble protein crystallization conditions. For example, many protein–protein complexes favor PEG as a precipitant over ammonium sulfate and it often takes less concentrated precipitant to crystallize the complex than required for its counterparts. When the known protein–protein complex crystallization conditions from this survey are grouped using a cluster analysis, excluding the oligomeric protein complex crystallizations, a set of 48 conditions is generated as the most probable conditions for screening of protein complex crystallizations. The conditions in the screen represent the centroids of each crystallization cluster. The majority of them are PEG conditions (39 PEG conditions *versus* 9 conditions involving ammonium sulfate and other salts) (Table 1). The crystallization conditions for oligomeric proteins, such as photoreaction centers, light-harvesting complexes, nucleosome and proteosome, appear to be unique to each individual complex and should be excluded from a general protein–protein complex screen.

This work is funded by the intramural research of the National Institute of Allergy and Infectious Diseases, National Institutes of Health.

References

Bergfors, T. M. (1999). Editor. *Protein Crystallization.* LaJolla, CA: International University Line.
Boyington, J. C., Motyka, S. A., Schuck, P., Brooks, A. G. & Sun, P. D. (2000). *Nature (London)*, **405**, 537–543.
Carter, C. W. Jr (1990). *Methods*, **1**, 12–24.
Carter, C. W. Jr & Carter, C. W. (1979). *J. Biol. Chem.* **254**, 12219–12223.
Garman, S. C., Wurzburg, B. A., Tarchevskaya, S. S., Kinet, J. P. & Jardetzky, T. S. (2000). *Nature (London)*, **406**, 259–266.
Gilliland, G. L. & Bickham, D. M. (1990). *Methods*, **1**, 6–11.
Hennessy, D., Buchanan, B., Subramanian, D., Wilkosz, P. A. & Rosenberg, J. M. (2000). *Acta Cryst.* D**56**, 817–827.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
McPherson, A. (1999). *Crystallization of Biological Macromolecules.* New York: Cold Spring Harbor Laboratory Press.
Radaev, S., Motyka, S., Fridman, W. H., Sautes-Fridman, C. & Sun, P. D. (2001). *J. Biol. Chem.* **276**, 16469–16477.
Radaev, S., Rostro, B., Brooks, A. G., Colonna, M. & Sun, P. D. (2001). *Immunity*, **15**, 1039–1049.